Biometrical Journal **52** (2010) 61, zzz–zzz / DOI: 10.1002/bimj.200100000

Quantifying intraclass correlations for nonnegative traits

Izabela R. C. Oliveira[*,1], **Geert Molenberghs** [2,3], **Clarice G. B. Demétrio** [4], **Carlos T. S. Dias** [4], and **Suely R. Giolo** [5]

[1] Department of Exact Sciences, Federal University of Lavras, 37200-000 Lavras, Brazil

[2] I-BioStat, Universiteit Hasselt, B-3500 Hasselt, Belgium

[3] I-BioStat, KU Leuven, B-3000 Leuven, Belgium

[4] Department of Exact Sciences, ESALQ/USP, 13418-900 Piracicaba, Brazil

[5] Department of Statistics, Federal University of Paraná, 80060-000 Curitiba, PR, Brazil

Received zzz, revised zzz, accepted zzz

The intraclass correlation is an important quantity in various areas of application. It is estimated based on fitting a model to hierarchical data and leads, in turn, to several concepts such as reliability, heritability, inter-rater agreement, etc. For data where linear models can be used, these attributes are conveniently defined as ratios of variance components. Matters are less simple for non-Gaussian outcomes. The focus here is on count and time-to-event outcomes where so-called combined models are used, extending generalized linear mixed models, to describe the data. These models combine normal and gamma random effects to allow for both correlation due to data hierarchies as well as overdispersion. Furthermore, because the models admit closed-form expressions for the mean, variances, higher moments, and even the joint marginal distribution, the derivation of intraclass correlations is convenient. The proposed methodology is illustrated using data from agricultural and livestock studies.

Key words: Generalized linear mixed model; Overdispersion; Poisson distribution; Random effect; Weibull distribution

Supporting Information for this article is available from the author or on the WWW under http://dx.doi.org/10.1022/bimj.XXXXXXX

1 Introduction

The intraclass correlation (ICC) coefficient has been formulated within the framework of random effects models and is frequently used to describe the degree of resemblance of units in the same cluster. The ICC is the foundation of such concepts as reliability and heritability; these are of great importance in a variety of applications (Bartlett and Frost , 2008; Hartl and Jones , 2001).

When the outcomes are, perhaps approximately, normally distributed, linear (mixed) models are frequently used (Verbeke and Molenberghs , 2000). The ICC then take the form of ratios of simple functions of variance components. The linear relationship between the response and the other model terms leads to simple, easily obtainable expressions.

However, when the trait of interest is not normally distributed and/or it does not follow a linear model, the random terms are no longer easily separable from the other model terms. This difficulty arises in particular when one deals with count and time-to-event outcomes, which are the focus of this paper. One often models such data using generalized linear models (McCullagh and Nelder , 1989), such as Poisson log-linear models for count data and Weibull, gamma, or exponential models for time-to-event data.

*Corresponding author: e-mail: izabela.rco@gmail.com, Phone: +55-35-8818-7966, Fax: +55-35-3829-1961

Table 1 *Tomatoes study. Descriptive statistics for glandular trichomes, considering 145 plants from F_2 population.*

Trichome	$\bar{x}$	s^2
Type IV	1.54	3.51
Type VI	0.06	0.41
Type VII	0.40	0.85

In empirical research, it has been observed recurrently that the mean-variance relationship for the Poisson model may not be met, with the same holding for time-to-event data. As a result, quite a bit of research has been devoted to formulate sufficiently flexible models for dealing with this phenomenon, referred to as over- or underdispersion. We will simply refer to it as overdispersion.

The so-called generalized linear mixed model (GLMM) (Breslow and Clayton , 1993; Molenberghs and Verbeke , 2005) has gained popularity in discrete hierarchical data modeling. When overdispersion and the need for hierarchical modeling occur simultaneously, the combined model family, proposed by Molenberghs *et al*. (2010) can be used. It accommodates overdispersion and clustering through two separate sets of random effects and contains as special cases the GLMM on the one hand, and several overdispersion models, such as the negative-binomial and beta-binomial models, on the other. In this paper, we make use of this family to handle overdispersion and correlated data, while obtaining ICCs based on nonnegative traits.

The proposed methodology will be illustrated using data from studies in agriculture and animal sciences. We will consider two case studies. In the first, different trichomes were counted on tomato plants and in the second, the time that progenies of Nelore sires, a Zebu breed, take to reach a pre-specified weight was recorded. For these data, we derive ICCs using Poisson and Weibull models.

The paper is organized as follows. In Section 2, the motivating cases are described with analyses reported in Section 5. A review of the Poisson and Weibull combined models for hierarchical and overdispersed data is the subject of Section 3. We use these combined models to obtain intraclass correlations for count and time-to-event traits, the expressions of which are presented in Section 4.

2 Case studies

We now describe two case studies, one featuring count data and the other giving rise to time-to-event data.

2.1 Inheritance study of trichomes density in tomato

An experiment was implemented at the Biology Department of Federal University of Lavras, Brazil, during the first semester of 2014 to study the inheritance of trichomes in tomato. Trichomes are epidermal outgrowths of diverse structure and function. The glandular ones are of particular importance since they secrete essences and urticant substances, which makes the plant resistant to some plagues of this crop.

A completely randomized designed experiment was implemented using plants from populations P_1, P_2, F_1, F_2, $BC_{1(1)}$ and $BC_{1(2)}$. Each experimental unit corresponded to one plant. In each plant, three cuts were made and at each of them an area of 1 mm^2 was defined. Then the quantities of several different trichomes were counted, both in the abaxial and adaxial faces of the leaves. It is reasonable to consider plant as a random effect.

Here, for illustrative purposes, we will consider only the glandular trichomes of types IV, VI, and VII counted in the adaxial face. Furthermore, we will use data from the F_2 population, which is the segregated one and consists of 145 plants. Table 1 presents descriptive statistics of the number of such trichomes in the F_2 population.

2.2 Nelore progenies data

These data were first analyzed in Giolo and Demétrio (2011). They refer to records from progeny of Nelore, a Zebu breed from India that has become predominant in Brazil. The original database consists of 3611 progenies of 24 Nelore sires and 3116 dams born during spring in a single herd between 1996 and 1997. About 16% of the dams had two progenies. For simplicity, we randomly selected one of the offspring of each of these dams. As a result, the data set used here consisted of 3116 progenies and the number of progenies per sire varied from 16 to 269. Evidently, a sire random effect will be needed.

All progenies were followed up from birth to approximately 2.5 years after they were born. Over this period, the weight of each progeny was taken six times at intervals of approximately 3–5 months, where the first weight was taken at birth. In addition to the weight, information about sex, reproduction (natural or artificial), progeny birth year, and age of the dam at progeny birth were recorded for each progeny.

Giolo and Demétrio (2011) compared sires based on length in days that their progenies need to gain 160 kg, a commercially specified weight gain employed in Brazil. Given that the weight is taken only periodically, the exact time that each progeny takes to gain the weight of interest was not precisely known. Thus, a logistic growth curve model for each progeny was fitted to estimate the exact response time. There are no censored data because all progenies have gained the targeted weight over the follow-up period.

In line with these authors, the estimated time to reach 160 kg is the response time of interest. Hierarchical models are sometimes difficult to fit, from a numerical standpoint, because the units used for outcomes and/or covariates lead to numerical values that are very large or small. So, not for statistical but rather for numerical reasons during model fitting, we use the months unit rather than the days one; evidently, this has no implications from a statistical view-point. Table 2 shows descriptive statistics for the estimated time, in months, that progenies took to gain 160 kg. For illustrative purposes we take the three sires whose progenies took the longest time and the three sires of which the progenies took the shortest time to gain this weight. More details about the data set and the estimation process are described in Giolo and Demétrio (2011).

Table 2 Descriptive statistics for estimated time (months) per sires that their progeny take to achieve the specified weight of 160 kg.

Mean	St. dev.	N	Min.	Max.
19.58	0.93	16	16.81	20.67
18.77	0.93	79	15.83	21.15
18.57	0.95	198	16.43	20.82
⋮	⋮	⋮	⋮	⋮
17.51	0.97	32	15.96	19.37
17.72	0.90	150	15.14	20.16
17.89	0.90	193	15.62	20.03

3 Extended models to handle hierarchical and overdispersed data

3.1 The Poisson combined model

In certain applications of standard generalized linear models, it is found that the data exhibit overdispersion, i.e., the variability is greater than predicted by the mean-variance relationship inherent in the model formulation. A number of models have been proposed for handling this phenomenon, especially in the Poisson case (Breslow , 1984; Lawless , 1987). Some also handle the rarer case of underdispersion.

An elegant formulation is through a two-stage approach. In the univariate Poisson case, we assume that $Y_i|\lambda_i \sim$Pois(λ_i) and then that λ_i is a random variable with $\mathrm{E}(\lambda_i) = \mu_i$ and $\mathrm{Var}(\lambda_i) = \sigma_i^2$. Then it follows that

$$\begin{aligned}\mathrm{E}(Y_i) =& \mathrm{E}\{\mathrm{E}(Y_i|\lambda_i)\} = \mathrm{E}(\lambda_i) = \mu_i,\\ \mathrm{Var}(Y_i) =& \mathrm{E}\{\mathrm{Var}(Y_i|\lambda_i)\} + \mathrm{Var}\{\mathrm{E}(Y_i|\lambda_i)\} = \mathrm{E}(\lambda_i) + \mathrm{Var}(\lambda_i) = \mu_i + \sigma_i^2.\end{aligned}$$

It is common to assume a gamma distribution for the random effects λ_i, leading to the so-called negative-binomial model (Hinde and Demétrio , 1998).

This model can be extended to the case of repeated measurements. We then assume a hierarchical data structure where Y_{ij} denotes the jth outcome measured for cluster i, $(i = 1, \ldots, N; j = 1, \ldots, n_i)$ and $\boldsymbol{Y}_i$ is the n_i-dimensional vector of all measurements available for cluster i. The scalar λ_i becomes a vector $\boldsymbol{\lambda}_i = (\lambda_{i1}, \ldots, \lambda_{in_i})'$, with $\mathrm{E}(\boldsymbol{\lambda}_i) = \boldsymbol{\mu}_i$ and $\mathrm{Var}(\boldsymbol{\lambda}_i) = \boldsymbol{\Sigma}_i$. Then, $\mathrm{E}(\boldsymbol{Y}_i) = \boldsymbol{\mu}_i$ and $\mathrm{Var}(\boldsymbol{Y}_i) = \boldsymbol{M}_i + \boldsymbol{\Sigma}_i$, where $\boldsymbol{M}_i$ is a diagonal matrix with the vector $\boldsymbol{\mu}_i$ along the main diagonal. The diagonal structure of $\boldsymbol{M}_i$ reflects the conditional independence assumption, that is, all dependence between measurements on the same unit stems from the random effects.

Alternatively, this repeated version of the overdispersion model can be combined with normal random effects in the linear predictor. Such models, proposed also by Dean (1991) and Thall and Vail (1990), for the count case, will be discussed next.

Molenberghs *et al.* (2007) specified a model for count data combining ideas from the overdispersion models and models with normal random effects. Later, Molenberghs *et al.* (2010) proposed a broad class of generalized linear models where the binary, count, and time-to-event cases were given particular emphasis. These models, named combined models, accommodate overdispersion and clustering through two separate sets of random effects and produce models with only normal random effects and models with only overdispersion terms as special cases.

Molenberghs *et al.* (2007) showed that the count models allow for closed-form expressions for the mean vector and variance-covariance matrix. As highlighted by the authors, the derivation of such closed forms has important implications because they admit, for example, explicit correlation expressions. This aspect was examined by Vangeneugden *et al.* (2011) for Poisson-type models for count data.

In line with Booth *et al.* (2003), Molenberghs *et al.* (2007) specified a model for repeated Poisson data with overdispersion:

$$\begin{aligned}Y_{ij} &\sim \mathrm{Poi}(\theta_{ij}\kappa_{ij}),\\ \kappa_{ij} &= \exp(x'_{ij}\boldsymbol{\beta} + z'_{ij}\boldsymbol{b}_i),\\ \boldsymbol{b}_i &\sim N(\mathbf{0}, D),\\ \mathrm{E}(\boldsymbol{\theta}_i) &= \mathrm{E}\{(\theta_{i1}, \ldots, \theta_{in_i})'\} = \boldsymbol{\Phi}_i,\\ \mathrm{Var}(\boldsymbol{\theta}_i) &= \Sigma_i,\end{aligned} \tag{1}$$

where Y_{ij} is the jth count outcome measured for cluster i, θ_{ij} is the random effect to capture overdispersion and κ_{ij} is the mean, which is linked to the linear predictor through the logarithmic link function. Notice that the normal random effect is placed in the linear predictor while the θ_{ij} random effect acts multiplicatively on the mean of the variable.

The θ_{ij} can be assumed to follow a gamma model, producing a so-called Poisson-Gamma-Normal (PGN) model or, equivalently, a negative-binomial-normal model. It is typically assumed that the components θ_{ij} of $\boldsymbol{\theta}_i$ are independent. In this case, Σ_i reduces to a diagonal matrix. It should be noted that it is possible to allow for general covariance structures; this is not considered further here.

Then, regarding the overdispersion random effects, three instances could be of interest: (1) the random-effects θ_{ij} are independent; (2) the θ_{ij} are allowed to be dependent; (3) they are equal to each other and hence reduce to $\theta_{ij} = \theta_i$. Independent θ_{ij} imply that the use is strictly confined to capture additional

overdispersion, i.e., not captured by the normal random effects. In contrast, when they are allowed to be correlated, they offer a way to model, for example, serial correlation.

The marginal mean vector and variance-covariance matrix were derived by Molenberghs *et al*. (2007) and are reproduced in Appendix A.1. We considered this model to calculate the between-cluster contribution to the total variability of the traits of interest, that is, the intraclass correlation. The derivation of such measure is presented in Section 4.

3.2 The Weibull combined model

Time-to-event data can be framed within the generalized linear model framework (McCullagh and Nelder , 1989). Next to this, the first of two aspects we need to take into account is overdispersion. The frailty model was introduced in survival analysis by Vaupel *et al*. (1979) as a way to allow unobserved heterogeneity. The model is an extension of the proportional hazards model in which the frailty term acts multiplicatively on the baseline hazard and captures the individual heterogeneity that refers to unobserved risk factors. It is assumed that this frailty term follows a parametric distribution; the most commonly used ones are gamma and log-normal. See also Duchateau and Janssen (2008). These models have also been extended to correlated time-to-event data, including in animal breeding studies (Schukken *et al*. , 2010; Giolo and Demétrio , 2011).

On the other hand, the generalized linear mixed model (GLMM; Breslow and Clayton (1993); Molenberghs and Verbeke (2005)) was formulated to handle hierarchical data of various outcome types, including time-to-event data. The model incorporates normally distributed random effects into the linear predictor of a generalized linear model.

Our modeling approach combines ideas from the GLMM and the gamma frailty model, in line with Molenberghs *et al*. (2010). As in the Poisson case, such models allow for closed-form expressions for the mean vector and variance-covariance matrix. The Weibull model for repeated measures, with both gamma and normal random effects, can be expressed as

$$\begin{aligned} f(\boldsymbol{y}_i|\boldsymbol{\theta}_i,\boldsymbol{b}_i) &= \prod_{j=1}^{n_i} \rho\theta_{ij} y_{ij}^{\rho-1} \exp(x'_{ij}\boldsymbol{\beta} + z'_{ij}\boldsymbol{b_i}) \exp\{-y_{ij}^{\rho}\theta_{ij} \exp(x'_{ij}\boldsymbol{\beta} + z'_{ij}\boldsymbol{b_i})\}, \\ \boldsymbol{b}_i &\sim N(\boldsymbol{0}, D), \\ \theta_{ij} &\sim \text{Gamma}(\alpha_j, \alpha_j^{-1}), \end{aligned} \tag{2}$$

where ρ is the Weibull shape parameter, $\boldsymbol{\beta}$ is a vector of fixed effects parameters, $\boldsymbol{b_i}$ is a vector of normal random effects with mean $\boldsymbol{0}$ and covariance matrix D. The covariate vectors x_{ij} and z_{ij} are for the fixed and random effects, respectively. Further, θ_{ij} are Gamma distributed. These effects are assumed independent while the correlation between the repeated measures is induced by normal random effects. Because of this, the normal random effects model the correlation between repeated measures and capture a certain amount of overdispersion, while the gamma random effects capture the remaining overdispersion. Should there be only normal random effects, the model might not be flexible enough to capture, say, a situation with weak correlation but strong overdispersion. The model is quite flexible from a distributional angle as well. For example, the Weibull distribution contains the exponential and gamma distributions as special cases. Admittedly, some distributional forms, such as the log-normal distribution, do not follow from this calculations and would have to be done separately.

Setting $\rho = 1$ leads to the special case of an exponential time-to event distribution. It is also evident that the classical gamma frailty model (no normal random effects) and the Weibull-based GLMM (no gamma random effects) follow as special cases.

The above expressions are derived for a two-parameter gamma density and because of the parametrization the mean of the gamma random effects is equal to 1. Also, while the α_j are allowed to depend on the measurement occasion or repetition within a subject, they can be held constant, as will be done in our case

study. Closed-forms expressions for the marginal density, means, variances, and covariances were derived by Molenberghs *et al.* (2010). The ones needed for our purposes, i.e., to derive closed-form intraclass correlation, are presented in Appendix A.2.

It is possible to also incorporate censored observations, as was done by Molenberghs *et al.* (2014). Once the model fitted to possible censored data, the computations outlined in the next section still apply. This is one of the several reasons to use methodology properly developed for time-to-event data, rather than applying a transformation.

4 Derivation of intraclass correlation

4.1 The count case

Consider the Poisson-Gamma-Normal model and its variance presented in (13). Also, without loss of generality, we set $\mathrm{E}(\boldsymbol{\theta}_i) = 1$. The variance is

$$\mathrm{Var}(Y_{ij}) = \mu_{ij} + \mu_{ij}(P_{i,jj} - 1)\mu_{ij}, \tag{3}$$

where

$$\mu_{ij} = \exp\left(x'_{ij}\boldsymbol{\beta} + \frac{1}{2}z'_{ij}Dz_{ij}\right) = \mu_{0ij}\mu_{1ij},$$

with notation as in (1), and

$$P_{i,jj} = \mu_{1ij}(\sigma_{i,jj} + 1)\mu_{1ij} = \mu_{1ij}^2(\sigma_{i,jj} + 1).$$

The contribution from overdispersion and from within-cluster variability over the total variability is:

$$\xi_{ij} = \frac{\mu_{0ij} + \mu_{0ij}\{(\sigma_{i,jj} + 1) - 1\}\mu_{0ij}}{\mu_{0ij}\mu_{1ij} + \mu_{0ij}\mu_{1ij}\{\mu_{1ij}^2(\sigma_{i,jj} + 1) - 1\}\mu_{0ij}\mu_{1ij}},$$

which reduces to

$$\xi_{ij} = \frac{1 + \mu_{0ij}\{(\sigma_{i,jj} + 1) - 1\}}{\mu_{1ij}[1 + \mu_{0ij}\mu_{1ij}\{\mu_{1ij}^2(\sigma_{i,jj} + 1) - 1\}]}. \tag{4}$$

These ratios place the variance presented in (3), for $D = 0$, in the numerator and the full variance of the combined model in the denominator. The proportion of the total variability that is between cluster, that is, the intraclass correlation is:

$$\mathrm{ICC}_{ij} = 1 - \xi_{ij}. \tag{5}$$

Because of the mean-variance relationship, ξ_{ij} and hence ICC_{ij} depends on the mean, therefore also on the covariates.

A specific case arises when there is no overdispersion. In such case, ξ_{ij} can be derived from the Poisson-Normal model or simply by setting $\sigma_{i,jj} = 0$ in (4). The contribution then is

$$\xi_{PNij} = \frac{1}{\mu_{1ij}\{1 + \mu_{0ij}\mu_{1ij}(\mu_{1ij}^2 - 1)\}}, \tag{6}$$

and the ICC_{ij} is calculated from replacing ξ_{ij} by ξ_{PNij} in (5).

It is not uncommon to model other sources of variability, by including other effect as a random term in model (1), which slightly changes the terms:

$$\begin{aligned}
\lambda_{ij} &= \exp(x'_{ij}\boldsymbol{\beta} + z'_{1ij}\boldsymbol{v}_i + z'_{2ij}\boldsymbol{w}_i),\\
\boldsymbol{v}_i &\sim N(\mathbf{0}, D_1),\\
\boldsymbol{w}_i &\sim N(\mathbf{0}, D_2).
\end{aligned}$$

These normal random effects are assumed to be independent, and its variance is:

$$\text{Var}(Y_{ij}) = \mu_{ij} + \mu_{ij}(P_{i,jj} - 1)\mu_{ij},$$

where

$$\mu_{ij} = \exp\left(x'_{ij}\beta + \frac{1}{2}z'_{1ij}D_1 z_{1ij} + \frac{1}{2}z'_{2ij}D_2 z_{2ij}\right) = \mu_{0ij}\mu_{1ij}\mu_{2ij}$$

and

$$P_{i,jj} = \mu_{1ij}\mu_{2ij}(\sigma_{i,jj} + 1)\mu_{1ij}\mu_{2ij} = \mu_{1ij}^2\mu_{2ij}^2(\sigma_{i,jj} + 1).$$

In this case, the contribution from overdispersion and from the random effect $\boldsymbol{w}_i$ over the total variability is:

$$\xi_{ij} = \frac{1 + \mu_{0ij}\mu_{2ij}\{\mu_{2ij}(\sigma_{i,jj} + 1) - 1\}}{\mu_{1ij}[1 + \mu_{0ij}\mu_{1ij}\mu_{2ij}\{\mu_{1ij}^2\mu_{2ij}^2(\sigma_{i,jj} + 1) - 1\}]}. \tag{7}$$

When there is no overdispersion, (7) reduces to:

$$\xi_{PNij} = \frac{1 + \mu_{0ij}\mu_{2ij}(\mu_{2ij} - 1)}{\mu_{1ij}\{1 + \mu_{0ij}\mu_{1ij}\mu_{2ij}(\mu_{1ij}^2\mu_{2ij}^2 - 1)\}}. \tag{8}$$

The intraclass correlations are obtained from applying (7) or (8) to (5).

Notice that the ratios ξ_{ij} and ξ_{PNij} are not free of the marginal mean function. In practice, therefore, one should compute some useful summaries of the values ICC_{ij}, given that they depend on the means at measurement j for unit i. Of course, when covariates are limited to a few factors with a limited number of levels, the intraclass correlation will only be dependent on these.

4.2 The time-to-event case

We here present expressions for the ICC in the Weibull and exponential cases by using the combined model, reviewed in Section 3. Details regarding calculations can be found in Appendix A.2. Further generalizations are available as web-based Supplementary Material.

Consider the combined model for overdispersion and correlated data presented in (2). Molenberghs *et al*. (2010) showed that its variance equals (14). The contribution from overdispersion and within cluster variability over the total variability is

$$\xi_{Wij} = \frac{\Gamma(\alpha_j)\Gamma\left(\alpha_j - \frac{2}{\rho}\right)\Gamma\left(\frac{2}{\rho} + 1\right) - \Gamma\left(\alpha_j - \frac{1}{\rho}\right)^2\Gamma\left(\frac{1}{\rho} + 1\right)^2}{\exp\left(\frac{1}{\rho^2}z'_{ij}Dz_{ij}\right)\zeta(\alpha_j, \rho, D, \boldsymbol{z}_{ij})}, \tag{9}$$

with

$$\begin{aligned}\zeta(\alpha_j, \rho, D, \boldsymbol{z}_{ij}) = &\left\{\Gamma(\alpha_j)\Gamma\left(\alpha_j - \frac{2}{\rho}\right)\Gamma\left(\frac{2}{\rho} + 1\right)\exp\left(\frac{1}{\rho^2}z'_{ij}Dz_{ij}\right)\right.\\ &\left.-\Gamma\left(\alpha_j - \frac{1}{\rho}\right)^2\Gamma\left(\frac{1}{\rho} + 1\right)^2\right\}.\end{aligned}$$

As in the Poisson case, the ratio in (9) places the variance of the combined model, considering $D = 0$, in the numerator and the full variance of the combined model in the denominator. The ICC is obtained by:

$$\text{ICC}_{Wij} = 1 - \xi_{Wij}, \tag{10}$$

that is the variability not related to the overdispersion or the proportion in the total variability due to the normal random term.

Very importantly, the above expressions do not depend on the fixed-effects structure. In other words, the only subject-specific influence comes from the random-effects design z_{ij}. This creates opportunities for constant intraclass correlations.

Setting $\rho = 1$ leads to the special case of an exponential time-to-event distribution. For this case the contribution from overdispersion and within cluster variability is

$$\xi_{Eij} = \frac{\alpha_j}{\exp(z'_{ij}Dz_{ij})\{2(\alpha_j - 1)\exp(z'_{ij}Dz_{ij}) - (\alpha_j - 2)\}},$$

and the intraclass correlation is obtained by $\text{ICC}_{Eij} = 1 - \xi_{Eij}$. Specific cases arise when there is only one variance component, d, in D. In this case, $\xi_{Wj} \equiv \xi_{Wij}$ reduces to

$$\xi_{Wj} = \frac{\Gamma(\alpha_j)\Gamma\left(\alpha_j - \frac{2}{\rho}\right)\Gamma\left(\frac{2}{\rho}+1\right) - \Gamma\left(\alpha_j - \frac{1}{\rho}\right)^2\Gamma\left(\frac{1}{\rho}+1\right)^2}{\exp\left(\frac{1}{\rho^2}d\right)\zeta(\alpha_j,\rho,d)}, \tag{11}$$

with

$$\zeta(\alpha_j,\rho,d) = \left\{\Gamma(\alpha_j)\Gamma\left(\alpha_j - \frac{2}{\rho}\right)\Gamma\left(\frac{2}{\rho}+1\right)\exp\left(\frac{1}{\rho^2}d\right) - \Gamma\left(\alpha_j - \frac{1}{\rho}\right)^2\Gamma\left(\frac{1}{\rho}+1\right)^2\right\}$$

and ξ_{Ej} reduces to

$$\xi_{Ej} = \frac{\alpha_j}{\exp(d)\{2(\alpha_j - 1)\exp(d) - (\alpha_j - 2)\}}.$$

Evidently, when the gamma parameter α_j is held constant, then $\xi_E \equiv \xi_{Ej}$, a constant.

As the Weibull-Normal model is a special case of the combined model, we can also obtain the intraclass correlation based on this model. For this, we derived the variance expression for this specific case. See Appendix A.3 for the derivation. The contribution from overdispersion and within cluster variability over the total variability, considering the Weibull-Normal model, is:

$$\xi_{WNij} = \frac{\xi_{N,WNij}}{\xi_{D,WNij}}, \tag{12}$$

where

$$\xi_{N,WNij} = \Gamma\left(\frac{2}{\rho}+1\right) - \Gamma\left(\frac{1}{\rho}+1\right)^2$$

and

$$\xi_{D,WNij} = \exp\left(\frac{1}{\rho^2}z'_{ij}Dz_{ij}\right)\left\{\Gamma\left(\frac{2}{\rho}+1\right)\exp\left(\frac{1}{\rho^2}z'_{ij}Dz_{ij}\right) - \Gamma\left(\frac{1}{\rho}+1\right)^2\right\}.$$

Considering the Exponential-Normal model, this contribution is:

$$\xi_{ENij} = \frac{1}{\exp(z'_{ij}Dz_{ij})\{2\exp(z'_{ij}Dz_{ij}) - 1\}}.$$

5 Analysis of case studies

5.1 Inheritance study of trichomes density in tomato

In this section, we will analyze the tomato data, introduced in Section 2.1. Our interest lies in measuring the magnitude of the measurement error to the variability between plants. This concept is known as reliability, a specific use of ICC. Here, we considered the Poisson combined model and its special cases to obtain such correlation.

Let Y_{ij} the number of trichomes counted in the j-th position of the i-th plant, where $i = 1, 2, \ldots, 145$ and $j = 1, 2, 3$, and the choose $\log(\kappa_{ij}) = \beta_0 + b_i$, where β_0 is the overall effect and b_i is the random effect that captures the plant-level variability, assumed to be normally distributed with mean 0 and variance d.

Results from fitting: (1) Poisson (P--) (model without random effects), (2) Poisson-Normal (P-N) (with the normal random effect), (3) Poisson-Gamma (PG-) (with the gamma random effect), and (4) Poisson-Gamma-Normal (PGN) (with both normal and gamma random effects) models are displayed in Table 3. Observe that, for all trichomes considered, the (P-N) is a strong improvement, in terms of likelihood, over the (P--) model. This is not surprising because correlations between the measures of the same plant are expected. Improvements in goodness-of-fit are also obtained when including the gamma random effect, especially using the combined model, which allows for correlation and overdispersion modeling. Note that in this case there is overdispersion caused by extra variability, which needs to be modeled. We then calculated intraclass correlation for each trichome type by applying the combined model (PGN) estimates in (4) and (5).

Given that there is only the plant random effect in this study, $z'_{ij}Dz_{ij}$ reduces to the scalar d. Also the θ_{ij} are assumed independent, with variance represented by $\sigma_{i,jj}$. Because we have no covariate effects, the mean value of the fixed structure, μ_{0ij} , can be obtained from only the intercept (overall) effect.

The contribution from model-induced measurement error and overdispersion to the total variability of this population is $\widehat{\xi} \cong 9.39 \times 10^{-6}$ for the trichome type IV, $\widehat{\xi} \cong 0.02$ for the trichome type VI and $\widehat{\xi} \cong 0.36$ for the trichome type VII. The intraclass correlations are $\widehat{\text{ICC}} \cong 0.99$ for trichome type IV, $\widehat{\text{ICC}} \cong 0.98$ for trichome type VI and $\widehat{\text{ICC}} \cong 0.64$ for trichome type VII. Such correlations can be interpreted as the proportion of the total variance that is due to the variability between plants. For the trichome type VII, for instance, this means that 64% of variability in measurements is due to differences in number of trichomes between plants, with the remaining 36% being due to overdispersion and measurement error. Notice that the ICCs (reliability) are higher for trichomes types IV and VI, which implies small measurements errors in comparison to the true differences between plants and relatively good distinction between these subjects.

5.2 Nelore progenies study

Giolo and Demétrio (2011) analyzed the data introduced in Section 2.2 using correlated frailty models with the baseline hazard function completely unspecified (Cox model). We will re-analyze the data but considering exponential and Weibull models and a predictor of the form:

$$\kappa_{ij} = \beta_0 + b_i + \beta_1 S_i + \beta_2 P_i,$$

where S_i is an indicator for progeny sex (1: male; 0: female), P_i is an indicator for progeny birth year (1: 1996; 0: 1997), and $b_i \sim N(0, d)$, the sire-level variance. We will use the intraclass correlation to quantify the degree of similarity between progenies from the same sire in terms of time, in months, they take to gain 160 kg. In quantitative genetics, this concept is known as heritability (Jacquard , 1983; Falconer and Mackay , 1996).

We consider special cases: (1) the purely exponential model (E--), (2) the purely Weibull model (W--), (3) the Weibull-Gamma model (WG-), (4) the Exponential-Normal model (E-N), (5) the Weibull-Normal model (W-N), and (6) the Weibull-Gamma-Normal model (WGN). Results from fitting all six models are displayed in Table 4. Clearly, the goodness-of-fit of the Weibull models is much higher than for the exponential models. This is not surprising, giving the high value of the estimated shape parameter $\widehat{\rho}$. Fitting

Table 3 *Tomatoes study. Parameter estimates and standard errors for the regression coefficients in (1) the purely Poisson model (P--), (2) the Poisson-Normal model (P-N), (3) the Poisson-Gamma model (PG-), and (4) the Poisson-Gamma-Normal model (PGN) for trichomes density, considering data from F_2 population.*

Effect	Par.	(P--)	(P-N)	(PG-)	(PGN)
			Trichome type IV		
Intercept	β_0	0.43 (0.04)	-1.72 (0.31)	0.43 (0.12)	-1.66 (0.31)
Gamma param.	α			0.17 (0.02)	3.15 (1.03)
Variance between plants	d		6.42(1.40)		6.31 (1.40)
-2log-likelihood		2440.1	1114.9	-436.8	-594.3
			Trichome type VI		
Intercept	β_0	-2.82 (0.20)	-4.84 (0.77)	-2.82 (0.29)	-4.36 (0.83)
Gamma param.	α			0.05 (0.02)	0.32 (0.28)
Variance between plants	d		4.37 (2.12)		3.34 (2.16)
-2log-likelihood		219.7	180.4	154.0	149.1
			Trichome type VII		
Intercept	β_0	-0.92 (0.08)	-1.42 (0.16)	-0.92 (0.10)	-1.24 (0.17)
Gamma param.	α			0.46 (0.10)	0.98 (0.41)
Variance between plants	d		1.06 (0.30)		0.70 (0.32)
-2log-likelihood		785.2	729.7	600.8	594.6

the Exponential-Gamma and Exponential combined models was not possible. This is not a total surprise. Molenberghs and Verbeke (2011) showed that the WGN model and its sub-models have a finite number of finite moments only. Precisely how many depends on the values of α and ρ: the order k of the moment should be $k < \rho\alpha$ for it to be finite. Given that the estimated ρ is high for the Weibull models, and the α's are in the order of magnitude of unity, there is no problem for the Weibull-based models considered here. However, setting $\rho = 1$ to obtain exponential versions, and assuming that α would be in the same order of magnitude, might imply that, for example, variances and perhaps means are not finite, underscoring difficulty of fit.

Turning again to the Weibull models, both the (WG-) and the (W-N) models are improvements, in terms of the likelihood, relative to the (W--) model. But the strongest improvement in fit occurs when considering the combined (WGN) model. As highlighted by Molenberghs *et al*. (2010), this strongly affects the point and precision estimates of key quantities such as slope difference and slope ratio. In our case, the differences between the fixed effects coefficients estimates of the (WGN) and those of the (W--) and (W-N) models are clearly noticeable. It is not uncommon to consider the (W-N) model, accommodating the correlation between progenies of the same sire, but according to the fit statistics the (WGN) is more appropriate.

Note that, while the point estimates for the Weibull shape parameters are high, the corresponding standard errors are small, providing further evidence for proper convergence. Further, note that for the (E-N) the variance of the normal random effect d is estimated on the boundary of the parameter space, effectively removing the parameter from the model. This explains why the (E--) and (E-N) fits are identical. A similar phenomenon does not occur in the Weibull cases.

Based on the model fits, we can now calculate ICCs. Using the one variance component in the random intercept for the sire effect, we apply expression (10) using (11). Note that throughout Section 4.2, the Gamma distribution parameter α, is represented with an index for the individual j. That said, here and in several other applications α is kept constant. Based on the estimates for the combined model (WGN), as laid out in Table 4, we obtain $\widehat{\xi_W} = 0.9230$ and:

$$\widehat{\text{ICC}} = 1 - 0.9230 = 0.077 \text{ (s.e. } 0.7943).$$

Table 4 Nelore study. Parameter estimates and standard errors for the regression coefficients in (1) the purely Exponential model (E--), (2) the purely Weibull model (W--), (3) the Weibull-Gamma model (WG-), (4) the Exponential-Normal model (E-N), (5) the Weibull-Normal model (W-N), and (6) the Weibull-Gamma-Normal model (WGN).

		Estimate (standard error)		
Effect	Par.	(E--)	(W--)	(WG-)
Intercept	β_0	-2.92 (0.03)	-13.72 (0.24)	-107.19 (2.89)
Sex	β_1	0.08 (0.04)	1.11 (0.04)	2.70 (0.10)
Prog. birth	β_2	-0.02 (0.04)	0.11 (0.04)	-0.65 (0.06)
Weibull shape	ρ		4.53 (0.08)	36.68 (1.00)
Gamma param.	α			1.53 (0.13)
-2log-likelihood		24314.0	15495.0	7230.9
Effect	Par.	(E-N)	(W-N)	(WGN)
Intercept	β_0	-2.92 (0.03)	-75.01 (0.99)	-110.78 (3.05)
Sex	β_1	0.08 (0.04)	1.87 (0.05)	2.81 (0.10)
Prog. birth	β_2	-0.02 (0.04)	-0.27 (0.11)	-0.65 (0.17)
Weibull shape	ρ		25.50 (0.33)	37.88 (1.05)
Gamma param.	α			1.52 (0.14)
Var. of sires	d	6.33×10^{-14}	0.09 (0.03)	0.21 (0.08)
-2log-likelihood		24314.0	7399.0	7100.6

This means low correlation between progenies from the same sire and/or low variability between sires. The low level of intraclass correlation is entirely natural for this setting, because there is high overdispersion, resulting from a relatively small α. This implies that there is a strong source of measurement error, inhibiting a high level of such correlation.

When this overdispersion is not taken into account, different estimates for intraclass correlation can be found. This is the case of considering the model (W-N) instead of (WGN). To calculate the intraclass correlation for this case we apply expression (10) using (12). Based on the estimates for the (W-N) we obtain $\widehat{\xi_{WN}} = 0.9451$ and:

$$\widehat{\text{ICC}} = 1 - 0.9451 = 0.0549 \text{ (s.e. } 0.0234).$$

Both estimates are similar against the background of a relatively large standard error. This is not surprising given the use of the delta method for a ratio. Details on precision estimation can be founded in supplementary material. Also, the estimates are close the boundary of the quantity's range, adding to the complexity to obtain a narrow standard error. Alternatively, Fieller's intervals could be calculated to gauge the estimates' precision.

6 Concluding remarks

In this paper, we have derived expressions for intraclass correlations, based on hierarchical count and time-to-event data. The focus was on the so-called combined models, which brings together generalized linear models with both normal and gamma random effects, thus accommodating correlation between repeated measures and overdispersion. Importantly, the standard generalized linear mixed models are special cases of the combined models, implying that the derivations reported here also apply to the commonly encountered GLMM cases. Further, as shown in Molenberghs *et al*. (2010), other GLMMs, for other data types and/or with other link functions, can be extended to a combined model formulation.

The combined models and its GLMM sub-model admit closed-form expressions for means, variances, and higher-order moments. As a result, variance ratios have explicit expressions too. The intraclass correlations are sufficiently simple and appealing, in particular in special cases, to be of practical value. Using real data, we illustrate how this coefficient is related to the important concepts of reliability and heritability.

We want to reiterate that, in the count models, intraclass correlation is a function rather than a constant. At first sight, this is a drawback. However, it is a consequence from the mean-variance relationship in the models considered. If the model fits the data well, it can also be claimed to be a feature of the data. Practically, intraclass correlation changes with the effects present in the predictor functions. Evidently, one can summarize the functions in a variety of ways, using averages, medians, quartiles, ranges, etc. In the time-to-event models, while, unsurprisingly, the correlation is not as simple as when outcomes are Gaussian, it does not depend on fixed effects. This is an appealing feature of this model.

All analysis were performed using the `SAS NLMIXED` procedure due to its flexibility to deal with non-normal responses and because it allows the user to specify the likelihood function to their own taste and needs. The `SAS` codes are available as supplementary material on the journal's web page.

Acknowledgements Geert Molenberghs gratefully acknowledges support from IAP research Network P7/06 of the Belgian Government (Belgian Science Policy). This work was partially supported by grants from CNPq, National Council for Scientific and Technological Development, Brazilian science funding agency. The authors thank professor dr. Wilson Roberto Maluf for having kindly provided the trichome data set used as illustration.

Conflict of Interest
The authors have declared no conflict of interest.

Appendix

A.1. Model elements for the Poisson combined model

The mean and variance expressions for the Poisson combined model (1) were presented by Molenberghs *et al*. (2007). The mean vector $\boldsymbol{\mu}_i = \mathrm{E}(\boldsymbol{Y}_i)$ has components:

$$\mu_{ij} = \phi_{ij} \exp\left(x'_{ij}\boldsymbol{\beta} + \frac{1}{2} z'_{ij} D z_{ij}\right)$$

and the variance-covariance matrix is given by

$$\mathrm{Var}(\boldsymbol{Y}_i) = \boldsymbol{M}_i + \boldsymbol{M}_i(\boldsymbol{P}_i - \boldsymbol{J}_{n_i})\boldsymbol{M}_i, \tag{13}$$

where $\boldsymbol{M}_i$ is a diagonal matrix with the vector $\boldsymbol{\mu}_i$ along the diagonal, $\boldsymbol{J}_{n_i}$ is a $n_i \times n_i$ dimensional matrix of ones and the $(j,k)^{th}$ element of $\boldsymbol{P}_i$ equals:

$$p_{i,jk} = \exp\left(\frac{1}{2} z'_{ij} D z_{ik}\right) \frac{\sigma_{i,jk} + \phi_{ij}\phi_{ik}}{\phi_{ij}\phi_{ik}} \exp\left(\frac{1}{2} z'_{ik} D z_{ij}\right).$$

Note that $\sigma_{i,jk}$ is the (j,k)th element of $\boldsymbol{\Sigma}_i$.

For univariate data with a single random θ_i and single normal random effect $b_i \sim N(0,d)$, the above expressions reduce to

$$\mu_{ij} = \phi_i \exp\left(x'_i\boldsymbol{\beta} + \frac{1}{2} d\right)$$

and

$$\mathrm{Var}(Y_i) = \mu_i + \mu_i^2 \left(\exp d - 1 + \frac{\sigma_i^2}{\phi_i^2} \exp d\right),$$

where ϕ_i and σ_i^2 are, respectively, the mean and variance of θ_i.

These expressions also produce their simplified counterparts for the special cases, including the Poisson-normal model and the Poisson model. For instance, when only normal random effects are present, the mean vector components slightly simplify:

$$\mu_{ij} = \exp\left(x'_{ij}\boldsymbol{\beta} + \frac{1}{2}z'_{ij}Dz_{ij}\right),$$

and the variance-covariance matrix is

$$\text{Var}(\boldsymbol{Y}_i) = M_i + M_i\{\exp(Z_iDZ'_i) - J_{n_i}\}M_i.$$

A.2. Model elements for the Weibull combined model

From the marginal density of the combined model specified by (2), Molenberghs *et al*. (2010) derived the moment and mean expressions:

$$\begin{aligned}
\text{E}(Y_{ij}^k) &= \frac{\alpha_j B(\alpha_j - k/\rho, k/\rho + 1)}{\lambda^{k/\rho}\alpha_j^{-k/\rho}} \exp\left(-\frac{k}{\rho}x_{ij}\boldsymbol{\beta} + \frac{k^2}{2\rho^2}z'_{ij}Dz_{ij}\right), \\
\text{E}(Y_{ij}) &= \frac{\alpha_j B(\alpha_j - 1/\rho, 1/\rho + 1)}{\lambda^{1/\rho}\alpha_j^{-1/\rho}} \exp\left(-\frac{1}{\rho}x'_{ij}\boldsymbol{\beta} + \frac{1}{2\rho^2}z'_{ij}Dz_{ij}\right),
\end{aligned}$$

which easily leads to the variance expression:

$$\begin{aligned}
\text{Var}(Y_{ij}) &= \frac{\alpha_j}{\lambda^{2/\rho}\alpha_j^{-2\rho}} \exp\left(-\frac{2}{\rho}x'_{ij}\boldsymbol{\beta} + \frac{1}{\rho^2}z'_{ij}Dz_{ij}\right)\left\{B(\alpha_j - 2/\rho; 2/\rho + 1)\exp\left(\frac{1}{\rho^2}z'_{ij}Dz_{ij}\right)\right. \\
&\quad \left. - \alpha_j B\left(\alpha_j - \frac{1}{\rho}; \frac{1}{\rho} + 1\right)^2\right\}. \qquad (14)
\end{aligned}$$

The exponential version follows from setting $\rho = 1$ in the above expressions.

A.3. Model elements for the Weibull-Normal model

We will derive the moment, mean and variance expressions for the specific case of no overdispersion and only one normal random effect, i.e., the Weibull-Normal model (W-N). It can be easily obtained from the Weibull-based GLMM density as described in the following. From

$$f(\boldsymbol{y}_i|\boldsymbol{b}_i) = \prod_{j=1}^{n_i} \lambda\rho y_{ij}^{\rho-1} \exp(x'_{ij}\boldsymbol{\beta} + z'_{ij}\boldsymbol{b_i}) \exp\{-\lambda y_{ij}^{\rho} \exp(x'_{ij}\boldsymbol{\beta} + z'_{ij}\boldsymbol{b_i})\}$$

let us derive the moments by

$$\text{E}(Y_{ij}^k|\boldsymbol{b}_i) = \lambda\rho \int y_{ij}^{k+\rho-1} \exp(x'_{ij}\boldsymbol{\beta} + z'_{ij}\boldsymbol{b_i}) \exp\{-\lambda y_{ij}^{\rho} \exp(x'_{ij}\boldsymbol{\beta} + z'_{ij}\boldsymbol{b_i})\} dy_{ij},$$

i.e., the k-th moment conditional upon $\boldsymbol{b}_i$. Then consider $\varphi_{ij} = \lambda \exp(x'_{ij}\boldsymbol{\beta} + z'_{ij}\boldsymbol{b_i})$, from which we deduce

$$\text{E}(Y_{ij}^k|\boldsymbol{b}_i) = \rho \int y_{ij}^{k+\rho-1}\varphi_{ij} \exp(-\varphi_{ij}y_{ij}^{\rho}) dy_{ij} = \int y_{ij}^{k} \exp(-\varphi_{ij}y_{ij}^{\rho}) d(\varphi_{ij}y_{ij}^{\rho}).$$

Writing $t = \varphi_{ij}y_{ij}^{\rho}$ we have $y_{ij}^k = t^{k/\rho}\varphi_{ij}^{-k/\rho}$ and

$$\text{E}(Y_{ij}^k|\boldsymbol{b}_i) = \int t^{k/\rho}\varphi_{ij}^{-k/\rho} \exp(-t) dt = \varphi_{ij}^{-k/\rho} \int_0^{\infty} t^{k/\rho} \exp(-t) dt = \varphi_{ij}^{-k/\rho}\Gamma\left(\frac{k}{\rho} + 1\right).$$

Then, the moment expression is

$$\mathrm{E}(Y_{ij}^k) = \frac{\Gamma\left(\frac{k}{\rho}+1\right)}{\lambda^{k/\rho}\exp(\mu_{ij}k/\rho)}\exp\{(k^2/2\rho^2)z'_{ij}Dz_{ij}\},$$

where μ_{ij} is the fixed part of the linear predictor. From this expression we have the first and second moments:

$$\begin{aligned}
\mathrm{E}(Y_{ij}) &= \frac{\Gamma\left(\frac{1}{\rho}+1\right)}{\lambda^{1/\rho}}\exp\left(-\frac{1}{\rho}x'_{ij}\boldsymbol{\beta}+\frac{1}{2\rho^2}z'_{ij}Dz_{ij}\right),\\
\mathrm{E}(Y_{ij}^2) &= \frac{\Gamma\left(\frac{2}{\rho}+1\right)}{\lambda^{2/\rho}}\exp\left(-\frac{2}{\rho}x'_{ij}\boldsymbol{\beta}+\frac{4}{2\rho^2}z'_{ij}Dz_{ij}\right),
\end{aligned}$$

which easily lead to the variance expression:

$$\begin{aligned}
\mathrm{Var}(Y_{ij}) &= \frac{1}{\lambda^{2/\rho}}\exp\left(\frac{-2}{\rho}x'_{ij}\boldsymbol{\beta}+\frac{1}{\rho^2}z'_{ij}Dz_{ij}\right)\\
&\quad\times\left\{\Gamma\left(\frac{2}{\rho}+1\right)\exp\left(\frac{1}{\rho^2}z'_{ij}Dz_{ij}\right)-\Gamma\left(\frac{1}{\rho}+1\right)^2\right\}.
\end{aligned}$$

References

Bartlett, J. W. and Frost, C. (2008). Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound in Obstetrics & Gynecology* **31**, 466–475.

Booth, J. G., Casella, G., Friedl, H. and Hobert, J. P. (2003). Negative binomial loglinear mixed models. *Statistical Modelling* **3**, 179–181.

Breslow, N. (1984). Extra-Poisson variation in log-linear models. *Applied Statistics* **33**, 38–44.

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of American Statistical Association* **88**, 9–25.

Dean, C. B. (1991). Estimating equations for mixed-Poisson models. In: Estimating Functions (ed. V.P. Godambe), Oxford University Press, Oxford, 35–46.

Duchateau, L. and Janssen, P. (2008). *The Frailty Model*. Springer, New York.

Falconer, D. S. and Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. 4th ed. Addison Wesley Longman, Harlow, England.

Giolo, S. R. and Demétrio, C. G. B. (2011). A frailty modeling approach for parental effects in animal breeding. *Journal of Applied Statistics* **38**, 619–629.

Hartl, D. L. and Jones, E. W. (2001). *Genetics: Analysis of Genes and Genomes*. Jones and Bartlett, Sudbury, Massachusetts.

Hinde, J. and Demétrio C. G. B. (1998). Overdispersion: Models and estimation. *Computational Statistics & Data Analysis* **27**, 151–170.

Jacquard, A. (1983). Heritability: one word, three concepts. *Biometrics* **39**, 465–477.

Lawless, J. (1987). Negative binomial and mixed Poisson regression. *The Canadian Journal of Statistics* **15**, 209–225.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. Jones and Chapman & Hall, London.

Molenberghs, G. and Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. Springer, New York.

Molenberghs, G. and Verbeke, G. (2011). On the weibull-gamma frailty model, its infinite moments, and its connection to generalized log-logistic, logistic, cauchy, and extreme-value distributions. *Journal of Statistical Planning and Inference* **141**, 861–868.

Molenberghs, G., Verbeke, G. and Demétrio, C. G. B. (2007). An extended random-effects approach to modeling repeated, overdispersed count data. *Lifetime Data Analysis* **13**, 513–531.

Molenberghs, G., Verbeke, G., Demétrio, C. G. B. and Vieira, A. (2010). A family of generalized linear models for repeated measures with normal and conjugate random effects. *Statistical Science* **25**, 325–347.

Molenberghs, G., Verbeke, G., Efendi, A., Braekers, R. and Demétrio, C. G. B. (2014). A combined gamma frailty and normal random-effects model for repeated, overdispersed time-to-event data. *Statistical Methods in Medical Research* **0**, 1–19.

Schukken, Y. H., Bar, D., Hertl, J. and Gröhn, Y. T. (2010). Correlated time to event data: Modeling repeated clinical mastitis data from dairy cattle in New York state. *Preventive Veterinary Medicine* **97**, 150–156.

Thall, P. F. and Vail, S. C. (1990). Some covariance models for longitudinal count data with overdispersion. *Biometrics* **46**, 657–671.

Vangeneugden, T., Molenberghs, G., Verbeke, G. and Demétrio, C. G. B. (2011). Marginal correlation from an extended random-effects model for repeated and overdispersed counts. *Journal of Applied Statistics* **38**, 215–232.

Vaupel, J. W., Manton, K. G., Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* **16**, 439–454.

Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer-Verlag, New York.

